

ALEXANDRA CARVALHO GUIMARÃES

**OS METADADOS E SEU USO NOS SITES DAS INSTITUIÇÕES DE
ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA**

**Monografia apresentada à disciplina
Projeto de Pesquisa em Informação
II, como requisito parcial à
conclusão do Curso de Gestão da
Informação, Departamento de
Ciência e Gestão da Informação,
Setor de Ciências Sociais
Aplicadas, Universidade Federal do
Paraná.**

Orientadora: Prof^a Lígia Kraemer

**CURITIBA
2004**

SUMÁRIO

LISTA DE ILUSTRAÇÕES	iii
RESUMO	iii
1 INTRODUÇÃO	1
2 OBJETIVOS	3
2.1 OBJETIVO GERAL	3
2.2 OBJETIVOS ESPECÍFICOS	3
3 PROCEDIMENTOS METODOLÓGICOS	4
4 LITERATURA PERTINENTE	6
4.1 TECNOLOGIA E TECNOLOGIA DA INFORMAÇÃO	7
4.2 INTELIGÊNCIA ORGANIZACIONAL	9
4.3 FERRAMENTAS DE BUSCA	12
4.3.1 Motores de busca	13
4.3.2 Diretórios	14
4.3.3 Metabuscadores	15
4.4 METADADOS	16
4.4.1 Conceito de Metadados	16
4.4.2 Função de Metadados	18
4.4.3 Padrões de Metadados	19
4.4.3.1 Padrão <i>Dublin Core</i>	20
4.4.4 Linguagens de marcação para descrição de metadados	22
4.4.4.1 <i>Hypertext Markup Language</i> (HTML)	22
4.4.4.2 <i>Extensible Mark-up Language</i> (XML)	24
5 ANÁLISE DOS DADOS	26
6 CONSIDERAÇÕES FINAIS	41
REFERÊNCIAS	43
APÊNDICE 1 – SITES QUE TRAZEM OS INDICADORES DAS FERRAMENTAS DE BUSCAS MAIS UTILIZADAS NO BRASIL	46
APÊNDICE 2 - E-MAIL ENVIADO ÀS FERRAMENTAS DE BUSCA: GOOGLE, CADÊ E ALTAVISTA	47
APÊNDICE 3 – LISTA DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA E SEUS ENDEREÇOS ELETRÔNICOS	48

LISTA DE ILUSTRAÇÕES

GRÁFICO 1 – INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM SITES – 2003.....	28
GRÁFICO 2 – SITES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM METADADOS – 2003	28
QUADRO 1 – EXISTÊNCIA DE SITES E METADADOS DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA NA INTERNET – 2003	27
QUADRO 2 – METADADOS HTML UTILIZADOS PELAS HOME PAGES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA – 2003	29
QUADRO 3 - IDENTIFICAÇÃO DAS LINGUAGENS DE MARCAÇÃO UTILIZADAS NOS SITE DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA – 2003	31
QUADRO 4 – INCIDÊNCIA DOS TERMOS NO METADADO “PALAVRA – CHAVE” DOS SITES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA – 2003	33
QUADRO 5 – POSIÇÃO DOS SITES DAS INSTIUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITBA NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” – 2004.....	34
QUADRO 6 – POSIÇÃO DOS SITES DAS INSTIUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITBA QUE NÃO POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” – 2004	35
QUADRO 7 – POSIÇÃO DOS SITES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” – 2004	37
QUADRO 8 – POSIÇÃO DOS SITES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE NÃO POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” – 2004	40

RESUMO

A pesquisa tem por objetivo investigar se a tecnologia de metadados, é um recurso que pode ajudar as instituições de ensino superior da rede particular de Curitiba, na sua representatividade digital na *Internet*. Explora fontes bibliográficas para construção de um referencial teórico, que aborda temas sobre tecnologia e tecnologia da informação, inteligência organizacional, ferramentas de busca e metadados. Por meio de um estudo exploratório faz-se um levantamento de dados feito em campo. Analisa os *sítes* de trinta e seis instituições e permite o conhecimento de como o metadado esta sendo utilizado nestas instituições. Conclui-se que o metadado deve ser mais estudado e difundido em língua portuguesa em função do resultado do levantamento bibliográfico desta pesquisa. Recomenda-se novos estudos e desenvolvimento de sistemas de busca inteligentes e integrados com outras tecnologias associadas.

1 INTRODUÇÃO

O conhecimento humano, que está em constante crescimento e desenvolvimento vem, ao longo dos anos, sendo explicitado e acumulado em diferentes suportes informacionais. Mas, em tempos em que a informação passa por uma reavaliação de paradigmas, o ser humano é incitado a buscar novas formas de geração, disponibilização e acesso ao potencial informacional produzido e estocado em diferentes bancos de dados.

A tecnologia da informação tem igualmente, possibilitado uma maior flexibilidade e rapidez no acesso às informações.

A evolução da utilização de tecnologias de informação ainda não foi suficiente, até o momento, para superar as questões referentes à recuperação e ao compartilhamento pela semântica da informação, o que é facilmente verificado em buscas na *Internet*.

Desde o início da *Internet*, houve a preocupação de se criarem ferramentas para a localização de seus recursos informacionais. Com o advento da *Web*, começaram a surgir as ferramentas específicas para pesquisas de suas páginas e hoje vemos que a análise e a indexação da informação são, em grande parte, operações automáticas que utilizam a técnica da extração de palavras para representar o conteúdo do texto analisado. Nesse contexto, emergiu a técnica de metadados.

Contudo, experiências em uma organização educacional, levaram a constatação do pouco uso do recurso de metadados na gestão dos *sites* das organizações da área, pressupondo-se que, apesar de uma grande quantidade de textos sobre este assunto, poucos profissionais que atuam nesta área fazem o uso desta técnica e sua aplicação.

Este fato motivou a realização de uma pesquisa sobre o assunto a qual foi caracterizada como exploratória pois, SANTOS (2002, p. 26), a define como a “aproximação de um tema, e visa criar maior familiaridade, em relação a um fenômeno”.

A realização de uma pesquisa científica sobre este assunto contribui para o crescimento teórico-conceitual na área de gestão de conteúdos digitais e possibilita o diagnóstico do gerenciamento desses conteúdos nas organizações, o que pode

contribuir com a documentação e organização de forma estruturada de seus dados para, assim, obterem vantagem competitiva no mundo digital.

A fundamentação teórica da pesquisa tem como apoio tanto aspectos contextuais sobre o tema como o próprio estudo do tema metadados, o que compõe uma seção denominada Literatura Pertinente.

Para enriquecer o tema, um levantamento dados sobre a utilização de metadados nos *sítes* das Instituições de Ensino Superior (IES) da rede particular de Curitiba, mostra a realidade encontrada no mundo digital, na seção Análise dos Dados.

Por fim, uma síntese conclusiva e recomendações compõem a última seção deste relatório de pesquisa.

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo geral desta pesquisa é investigar se a tecnologia de metadados é um recurso que pode ajudar as IES da rede particular de Curitiba, na sua representatividade digital na *Internet*.

2.2 OBJETIVOS ESPECÍFICOS

Para atingir o objetivo geral foi necessário cumprir com os seguintes objetivos específicos:

- a) fazer um estudo teórico básico sobre metadados;
- b) mapear os *sítes* das IES da rede particular de Curitiba de acordo com o padrão de metadados utilizados;
- c) investigar qual a linguagem que estas organizações estão empregando na construção de seus *sítes*;
- d) identificar a política de indexação e recuperação dos buscadores da *Internet*;
- e) verificar, por meio dos buscadores, a posição que os *sítes* das IES da rede particular de Curitiba ocupam nos resultados das buscas frente ao uso ou não de metadados.

3 PROCEDIMENTOS METODOLÓGICOS

Para o desenvolvimento deste trabalho foi necessário seguir três etapas distintas: Literatura Pertinente, Análise dos Dados e Considerações Finais.

Na etapa Literatura Pertinente, fez-se um estudo teórico básico sobre tecnologia e tecnologia da informação, inteligência organizacional, ferramentas de busca e metadados. O levantamento bibliográfico foi realizado nas bibliotecas da Universidade Federal do Paraná (UFPR), da Pontifícia Universidade Católica do Paraná (PUC – PR) e da Celepar e também na *Internet*. O número de documentos encontrados foi elevado, porém sobre o contexto do uso dos metadados foi encontrada pouca literatura em língua portuguesa.

Na etapa Análise dos Dados, estão descritas as buscas feitas na *Internet* sobre os *sítes* das IES da rede particular de Curitiba.

A princípio, a pesquisa incluiria também a rede particular de ensino fundamental e médio da cidade de Curitiba, mas, no decorrer do levantamento das instituições constatou-se que a fonte mais precisa para obter dados fundamentais sobre estes níveis de ensino seria o Sindicato das Escolas Particulares o qual, impôs um custo para a obtenção do relatório necessário. Barreiras de cunho monetário impediram à princípio a inclusão das instituições desta rede na pesquisa.

Recorreu-se então, à Secretaria Municipal de Educação, a qual alegou que as escolas particulares não permitem o acesso aos dados pretendidos, o que impossibilitou novamente a inclusão daquelas instituições.

Foi, então, que se decidiu pelas Instituições de Ensino Superior, encontradas no *site* do Instituto Nacional de Ensino Superior (Inep), órgão que se responsabiliza pela veracidade de competência das instituições reconhecidas pelo Ministério da Educação e Cultura (MEC).

O período pesquisado no *site* do Inep abrangeu os meses de setembro a novembro de 2003 e o número de instituições identificadas até esta data, foi de 39 (trinta e nove).

O passo seguinte foi um levantamento de quais instituições que possuem *sítes*, quais utilizam padrões de metadados e que tipo de linguagem de marcação fazem uso, por meio das *home pages* de cada instituição,.

Em seguida foram consultadas as ferramentas de buscas da *Internet* mais utilizadas no Brasil, por meio de *sites* que trazem estes indicadores (apêndice 1), na tentativa de descobrir qual a política de indexação e recuperação de cada uma delas. No entanto, pela falta de informação desta natureza nos próprios *sites* das ferramentas, foram enviados *e-mails* solicitando tais informações (apêndice 2) , mas nenhuma resposta foi obtida.

Além disso, foram efetuadas buscas utilizando-se do metadado “palavras-chave” dos *sites* das IES da rede particular de Curitiba nas ferramentas de busca mais utilizadas, com a finalidade de identificar qual a classificação de cada *site* nos resultados de buscas.

Os dados levantados foram, então, tabulados e analisados, resultando em considerações finais e recomendações ao tema.

4 LITERATURA PERTINENTE

Com a evolução e a diminuição do custo dos recursos de Informática, tanto em *hardware* quanto em *software*, as empresas foram se beneficiando das tecnologias da informação para uma maior abrangência de sua atuação nos negócios das empresas.

As mudanças ambientais, ocorridas nos últimos tempos, impulsionadas principalmente pelo fator tecnológico, vêm propiciando e, ao mesmo tempo, forçando os serviços de informação a planejarem e redimensionarem os produtos e serviços oferecidos aos seus usuários. Dessa forma, os serviços tradicionais gradativamente estão sendo incorporados às Tecnologias da Informação.

O mundo está submetido a uma profunda mudança que talvez indique uma nova idade na história da civilização, uma mudança de conseqüências tão notáveis quanto às provocadas pelas grandes descobertas e pela Revolução Industrial que deram origem à Idade Moderna.

As mudanças estão sendo chamadas, apropriadamente, de Revolução da Tecnologia da Informação ou Revolução Informacional. Essa revolução se exprime, de um lado, por um novo paradigma tecnológico que se organiza em torno da Tecnologia da Informação e, de outro, por uma nova dimensão sócio-econômica potencializada pela estruturação global.

Diante destes fatos, dedica-se esta seção à compilação contextual de conteúdos referentes à Tecnologia, Tecnologia da Informação, Inteligência Organizacional, Ferramentas de Busca e Metadados na intenção de fundamentar teoricamente o estudo e proporcionar um maior entendimento sobre este assunto ao leitor.

4.1 TECNOLOGIA E TECNOLOGIA DA INFORMAÇÃO

De acordo com o dicionário *Michaelis* (2000, p. 580), tecnologia é definida como “aplicação dos conhecimentos científicos à produção em geral”. De forma mais abrangente, FLEURY (1993, p. 35) conceitua que a “tecnologia é um pacote de informações organizadas de diferentes tipos provenientes de várias fontes, obtidas por diferentes métodos e utilizada na produção de bens e serviços”. O autor ainda acrescenta que os conhecimentos e habilidades empregadas na produção desses pacotes tecnológicos constituem-se na capacitação tecnológica, a qual não se compra, mas se constrói ao longo do tempo, fruto de um processo evolutivo das pessoas e organizações.

Para PRETTO (1996, p.15) a tecnologia é um conjunto de mecanismos para aumentar as possibilidades de locomoção, produção e comunicação, ou seja, a tecnologia é uma extensão da capacidade humana.

Pode-se, então, definir tecnologia como um conjunto de conhecimentos especiais e principalmente científicos, que se aplica a um determinado ramo de atividade, podendo ser considerada também como uma ciência que trata da técnica.

Tecnologia da informação é a expressão usada para o conjunto dos conhecimentos que se aplicam em utilizações específicas, esta expressão trata de relações complexas entre sistemas de informação, uso e inovação de *hardware*, *software*, serviços e usuários.

Segundo KEEN (1996, p. 273), a tecnologia da informação é composta dos seguintes elementos: redes de comunicação, *links* de telecomunicações, *hardware*, *software* e depósitos de informação. A integração e evolução desses elementos revolucionam as formas de comunicação, as formas de pensar e de fazer negócios.

Para CASTELLS (2000, p. 58), as tecnologias da informação já eram observadas antes de 1940, mas foi durante a Segunda Guerra Mundial e no período seguiu que surgiram as tecnologias em eletrônica. Porém, a difusão em escala comercial dessas tecnologias só veio a ocorrer, verdadeiramente, a partir de meados dos anos 70 do século XX.

Na medida em que o uso comercial da tecnologia da informação foi se desenvolvendo e tornando-se mais sofisticado, ficou claro que os sistemas gerenciais deveriam se comunicar entre si. De uma maneira simplificada, isso

significa que interfaces deveriam ser construídas para permitir a troca de mensagens estruturadas entre os sistemas, o que requer a garantia de definições consistentes para os dados dentro de cada sistema, o que é freqüentemente ignorado.

No início dos anos 80 do século XX, o termo computador cobria todo o aspecto do processamento da informação. De acordo com KEEN (1996, p. 273), hoje a tecnologia da informação tornou-se a designação mais utilizada para uma gama crescente de equipamentos, aplicações, serviços e tecnologias básicas que se enquadram em três categorias: computadores, telecomunicações e dados de multimídia, cada um com suas subcategorias. O autor ainda enfatiza que a tecnologia da informação vive em contínuo crescimento, pois, todos os dias surgem inovações importantes.

Segundo GRAEML (2000, p. 17), as Telecomunicações e a Informática têm contribuindo para grandes transformações em nossa sociedade, ou seja, elas encurtam distâncias e permitem que máquinas assumam e executem com competência tarefas que exigiam muito esforço e tempo dos humanos.

FERRANTE e RODRIGUEZ (1995, p. 46) já afirmavam que “a melhoria da competitividade global principalmente em áreas não diretamente relacionadas à informação, ocorrerá com o uso correto de modernas tecnologias de informação”.

De acordo com GRAEML (2000, p. 69), a tecnologia da informação contribui para que a organização seja capaz de oferecer um produto ou serviço que a concorrência não terá condições de copiar rapidamente, principalmente se conseguir agregar a tecnologia a outros fatores de competitividade intrínsecos da empresa. O autor ainda acrescenta que a tecnologia da informação torna-se um importante ativo para alavancar as atividades, proporcionando vantagem competitiva e diferenciação dos negócios, principalmente relacionada a aplicações estratégicas.

As novas tecnologias da informação dispõem de grande capacidade de auxiliar na obtenção dos objetivos, podendo tornar-se, em alguns casos, o principal fator determinante do sucesso da empresa. De acordo com FERRANTE e RODRIGUEZ (1995, p. 45), “uma organização inserida na sociedade da informação deverá tirar total vantagem do uso de modernas tecnologias de informação, para ganhar competitividade”. Os autores há anos já previam que a tecnologia deveria abandonar seu papel secundário, de suporte, e se tornar um componente fundamental para a organização atingir qualidade e competitividade.

MARTIN (1991, p.47) descreve que um sistema totalmente informatizado, funcionando eficientemente e eficazmente, proporciona vantagens em relação ao tempo, à organização, à facilidade de obtenção de informações, à previsão e a muitos outros aspectos que contribuem para o sucesso da empresa.

KEEN (1996, p.17) já defendia a idéia de que a tecnologia da informação era uma força decisiva, ela representava a metade dos gastos do capital das organizações e, crescentemente, afetava como as empresas se organizavam, operavam e concorriam.

Conclui-se, portanto, que a tecnologia da informação pode ser entendida como um conjunto de meios utilizados para alavancar e potencializar o processo de criação e desenvolvimento de capacitação tecnológica, proporcionando vantagens competitivas. Sendo assim, pode-se afirmar que a implementação e o uso correto de tecnologias adequadas, conduz à melhora da competitividade global das organizações.

Em um nível maior de contexto, a tecnologia e a tecnologia da informação devem estar incorporadas à inteligência organizacional das empresas para aumentar sua competitividade no mercado global, e é sobre este tema que trata a seção seguinte.

4.2 INTELIGÊNCIA ORGANIZACIONAL

A inteligência organizacional é um ciclo contínuo de atividades que incluem o sensoramento do ambiente, o desenvolvimento de percepções e a criação de significados por intermédio de interpretação, utilizando a memória de acontecimentos passados e escolhendo ações baseadas nas interpretações desenvolvidas. De acordo com CHOO (1998, p. 20), o sensoramento se refere à busca de informações sobre o ambiente interno e externo.

Segundo MORESI (2001, p. 35), a inteligência organizacional apóia a criação do conhecimento organizacional que, potencializado pelo aprendizado organizacional, cria condições necessárias para que as organizações possam inovar e se adaptar à dinâmica e complexidade das condições ambientais.

De acordo SILVA (2003, p. 117), um novo paradigma da inteligência organizacional é a competitividade, ou seja, é a habilidade estratégica que todas as

organizações, das pequenas às grandes, com fins lucrativos ou sociais, devem adquirir para se manterem “alertas” e atualizadas, em um processo de monitoramento permanente. A autora afirma que a inteligência organizacional é um ciclo de análises que leva ao conhecimento para a tomada de decisão, o que pode implicar mudança total da estratégia, ou mesmo a redefinição dos negócios da organização.

Segundo ANSOFF (1990, p.96), “através da inteligência organizacional é possível estabelecer estratégias e objetivos, tanto em momentos diferentes quanto em níveis diversos de uma organização, assim atributos de desempenho podem ser um objetivo da empresa num momento e também podem ser como sua estratégia em outro momento”.

De acordo com BEUREN (2000, p.43), a inteligência organizacional passou a contemplar a estrutura das organizações pois, no mercado altamente competitivo, há necessidade de se direcionar os esforços tanto para a estabilidade quanto para as mudanças, ou seja, deve-se buscar ganhos de eficiência operacional, bem como manter sintonia com o ambiente externo.

A inteligência organizacional pode auxiliar as organizações de várias formas, entre elas CIANCONI (2001, p.54) destaca: na elaboração de decisões e estratégias, identificação de concorrentes no mercado, conhecimento de seus pontos fortes e fracos e, conseqüentemente, identificação de desafios, riscos e oportunidades de negócios. A autora acrescenta que a “inteligência organizacional consiste, portanto, em uma forma específica de gestão da informação, que coloca ênfase no ambiente competitivo e em conhecer os aspectos que podem causar impacto nos pontos fortes e fracos das organizações”.

As organizações necessitam de informações — informações de conteúdos específicos, disponibilizadas no momento certo, adaptadas de uma melhor forma e devidamente contextualizadas — para atingir competitividade no mercado. Para isso, é necessário o uso de tecnologias voltadas ao trabalho com informação que se concretizam basicamente em sistemas de informações adaptados para cada tipo de negócio.

Segundo MORAVEC (1965, p.37), um sistema de informação pode ser definido como os procedimentos, as metodologias, a organização, os elementos de

software e de *hardware* necessários para inserir e recuperar dados selecionados, conforme forem sendo exigidos para a operação e a gestão de uma organização.

Para GORDON (1974, p.25), um sistema de informação é um recurso integrado entre homem e máquina para fornecer informações em apoio às funções de operações, de administração e de tomada de decisões da organização utilizando *hardware* e *software* de computação, procedimento manuais, modelos de decisão administrativa e uma base de dados.

EIN-DOR (1985, p. 30), já afirmava que os sistemas de informações em geral têm como principal função coletar, armazenar, processar e produzir informações. O autor ainda assegura que no ambiente moderno, esses sistemas têm uma base de computação, incluindo o uso de modelos analíticos de decisão, bases de dados de larga escala e tecnologias de tempo compartilhado.

De acordo com MARTIN (1991, p.140), “os efeitos das ferramentas de produtividade de sistemas de informações, mostram que o resultado é melhor em algumas empresas e pior em outras, que utilizam as mesmas ferramentas”. Com a afirmação do autor, pode-se concluir que para conseguir uma produtividade alta, é necessário não apenas selecionar as melhores ferramentas, mas também adaptar a organização e os métodos a sistemas de informações voltados às suas necessidades.

Para BEUREN (2000, p. 36), adotar um sistema de informação adequado pode proporcionar à organização vantagens competitivas e permitir a visualização da organização em termos objetivos e funcionais.

Sistemas de informações devem fornecer informações passadas, presentes e projetadas, relacionadas com inteligências internas e externas e dar apoio às funções de planejamento e controle de uma organização dentro de um tempo adequado para a tomada de decisões acertadas, para assim inserir a organização no mercado competitivo revelando a sua inteligência.

A inteligência organizacional também está presente no mundo virtual. Muitas organizações utilizam a *Internet*, por meio de seus *sítes*, para se fazer representar no mercado virtual como a forma estratégica e competitiva para adquirir novos clientes. Para que isso ocorra, é necessário que as organizações, que estão inseridas neste mercado virtual, apareçam e estejam bem representadas na *Internet*.

Para serem encontradas no mundo virtual se faz necessário que as ferramentas de buscas as localizem e reconheçam, que é a abordagem da seção seguinte.

4.3 FERRAMENTAS DE BUSCA

Existem, hoje, centenas de ferramentas para busca, ou buscadores, de informações nas cerca de um bilhão de páginas digitais que se estimam existir na *Internet*. As peculiaridades destas ferramentas influenciam no tipo, número e qualidade dos recursos recuperados através delas. Em função disto apresenta-se aqui uma visão das principais categorias de ferramentas de busca da *Internet*, suas semelhanças, diferenças e características, bem como analisam-se as vantagens e desvantagens associadas a cada uma.

Uma ferramenta de busca para MOURA (2003) é um conjunto organizado constituídos de tecnologias apropriadas, índices, bases de dados e algoritmos, com o objetivo de analisar e indexar páginas da rede e armazená-las em uma base de dados.

Segundo CENDÓN (2003, p.12), desde os primórdios da Internet, houve a preocupação de se criarem ferramentas para a localização de seus recursos informacionais. Entre as ferramentas mais antigas podem ser citadas o *Archie*, que busca arquivos em repositórios de FTP, e a *Veronica* e a *Jughead*, que encontram conteúdos armazenados nos *Gophers*.

Com a grande quantidade de informações na *Web* e a conseqüente explosão das páginas digitais para pesquisa de seus conteúdos, houve a necessidade do desenvolvimento de sistemas que pudessem armazenar estas informações e também ter a possibilidade ao seu acesso por meio de buscas. Como conseqüência disto começaram a surgir dois tipos básicos de ferramentas de busca na *Internet*: os motores de busca, diretórios e os metabuscadores. Entretanto, a partir dessas duas categorias básicas, outros tipos de ferramentas têm surgido, fazendo com que o mundo dos serviços de busca se torne complexo e volátil.

As características específicas de cada ferramenta, o tipo, a quantidade e a qualidade dos recursos recuperados, podem variar imensamente.

Para que se obtenha melhores resultados nas buscas de informações, deve-se entender as particularidade dos diferentes tipos de ferramentas de busca na *Web*.

4.3.1 Motores de busca

Um motor de busca é tipo específico de ferramenta de busca, assim como diretórios e metamotores. No, entanto ele se preocupa menos com a seletividade do que com a abrangência de suas bases de dados, procurando colecionar o maior número possível de recursos, fazendo uso de programas chamados de robôs de busca.

De acordo com a MOURA (2003) motores de buscas são os sistemas baseados no uso exclusivo de programas de computador para a indexação e recuperação das páginas da *Web*.

Segundo CENDÓN (2001, p. 41), “os motores de busca começaram a surgir quando o número de recursos na *Web* adquiriu proporções tais que impediam a sua coleta por meios manuais e a busca apenas através da navegação”. A autora ainda traz a informação de que o primeiro motor de busca foi o *WebCrawler*, lançado em abril de 1994, o qual deu origem a vários tipos de motores, contudo, estes foram desativados na medida em que a tarefa a ser executada passou a exigir maiores recursos humanos e técnicos.

Os robôs de busca denomina-se também de aranhas (*spiders*), agentes, viajantes (*wanderers*) e rastejadores (*crawlers*), que são programas lançados na *Internet* que tem como principal função obter dados sobre o maior número possível de documentos para integrá-los, posteriormente, à sua base de dados. Os motores de busca podem usar vários robôs para trabalharem em paralelo.

Suas principais características são suas enormes bases de dados, podendo alcançar milhões de itens, permitem aos usuários localizar os itens desejados mediante buscas por palavras-chave.

Atualmente os motores de busca utilizam-se de quatro componentes: um robô, que localiza e busca documentos na *Internet*; um indexador, que extrai a informação dos documentos e alimenta em base de dados; o motor de busca propriamente dito; e a interface, que é utilizada pelos usuários.

Segundo BACCHIN (2003), os robôs visitam automaticamente as páginas da Internet, e as incluem em seus bancos de dados. Porém, uma página aceita e incluída no banco de dados automaticamente, depende da forma como ela foi feita. Os robôs vasculham a *Internet* e vão adicionando as páginas no banco de dados. Se for alterada a página, *layout* ou conteúdo, e os robôs visitarem sua página depois destas mudanças, isto pode e provavelmente vai afetar seu posicionamento nas buscas.

De acordo CENDÓN (2001, p.41) “os documentos encontrados pelos robôs são encaminhados aos indexadores que extraem as informações das páginas HTML e as armazenam em uma base de dados”.

Os motores de busca diferem entre si em relação ao tamanho de suas bases de dados, os critérios para indexação e inclusão de páginas, a aplicação de seus resultados, suas interfaces, os recursos de busca que oferecem e a frequência com que atualizam suas bases de dados

As bases de dados de cada motor não são iguais. Ou seja, para a mesma busca, cada mecanismo invariavelmente trará resultados que outros não encontraram. Assim, para realização de uma busca recomenda-se que seja utilizada mais de uma ferramenta.

4.3.2 Diretórios

O diretório foi a primeira solução indicada para organizar e localizar os recursos da *Web*, tendo precedido os motores de busca por palavras-chave. Segundo CEDÓN (2001, p. 39), os diretórios “foram introduzidos quando o conteúdo da *Internet* ainda era pequeno o suficiente para permitir que fosse coletado de forma não automática”.

De acordo com BACCHIN (2003) “os diretórios são sistemas de busca controlados por editores humanos”, ou seja, o processo de cadastro de *sites* passa por uma verificação não-automática.

Os diretórios constituem-se em bancos de dados que estão divididos em categorias e subdivididos em subcategorias, ou seja, os *sites* recebem uma organização hierárquica de assunto e permitem aos usuários localizar informações, navegando, progressivamente, para as subcategorias. Sendo assim, os cabeçalhos

de assunto são atribuídos de forma consistente, de modo que os usuários podem contar com a ajuda de um vocabulário controlado. Como são ferramentas destinadas a um público variado, procuram incluir, em suas árvores hierárquicas de assuntos, tópicos que são de interesse amplo.

Segundo CENDÓN (2001, p.40), existem, por exemplo, diretórios especializados em imagens, jornais e revistas, *software*, listas de discussão; outros coletam *sites* sobre assuntos específicos como saúde, ciências, legislação, informática etc.; ainda outros listam ferramentas de busca de países específicos ou para um público-alvo determinado.

Com a frequência com que novas ferramentas de busca surgem, ao mesmo tempo em que outras caem em desuso, um novo tipo diretório passou a ser criado — o diretório de ferramentas de busca que têm o objetivo de listar buscadores para facilitar a identificação da própria ferramenta. Exemplo deste tipo de diretório é o *FinderSeeker* ou *Search.com*.

4.3.3 Metabuscadores

Os metabuscadores são ferramentas de busca que se utilizam das bases de dados tanto dos motores de buscas quanto dos diretórios para realizar a pesquisa a *Web*.

Os metabuscadores não possuem uma base de dados própria, por isso procuram em várias ferramentas de busca ao mesmo tempo.

De acordo com CENDÓN (2001, p.47), metabuscadores são ferramentas que não possuem base de dados, ou seja, ele utiliza dados de outras ferramentas de busca. Exemplos de metamotores são *Inference Find*, *SavvySearch*, *Mamma*, *MetaMiner* e *MetaBusca ZAZ*, sendo estes dois últimos brasileiros.

Existem variações entre os metabuscadores. Eles apresentam diferenças em relação à interface de busca, motores utilizados na pesquisa, modo de processamento das consultas, bem como forma de compilação e apresentação dos resultados.

Segundo BACCHIN (2003), os metabuscadores podem funcionar como os diretórios de ferramentas temáticas, permitindo que os usuários selecionem ferramentas especializadas em algum idioma ou assunto.

Os metabuscadores também podem efetuar buscas em outras partes da *Internet* como os arquivos de grupos de discussão da *Usenet* ou em *newswires*.

Quanto ao processamento da consulta, a maioria dos metabuscadores permite a formulação de uma expressão de busca em uma sintaxe semelhante à usada pela maioria dos motores, podendo permitir também o uso de lógica booleana e mesmo de linguagem natural.

A tecnologia metadados tem como uma de suas funções auxiliar as ferramentas de buscas a obterem maior precisão nos resultados em suas buscas. Na próxima seção serão apontadas algumas características desta tecnologia.

4.4 METADADOS

Metadado é um recurso aplicado à informação e para compreender seu potencial e função, seguem considerações e características a seu respeito

4.4.1 Conceito de Metadados

Metadados é um assunto que vem sendo tratado há muito tempo em sistemas de processamento da informação. O que tem mudado é o escopo de sua atuação conforme a evolução da tecnologia da informação.

As pessoas que têm contato com ferramentas de tecnologias da informação utilizam alguma forma de metadados, mesmo eventualmente sem saber que a usam pois há controvérsias sobre sua utilização.

A definição mais comum para metadados é “dados sobre dados”. Embora essa definição seja uma expressão da verdade e englobe todas as tentativas de definição, ela não dá a verdadeira dimensão de todos os aspectos que o termo envolve.

KEEN (1996, p. 273) afirma que os metadados descrevem atributos de um recurso, suportando um número de funções como: localização, descoberta, avaliação, seleção, entre outros. O autor ainda afirma que o recurso de metadados é uma tecnologia da informação que está associada a um dado com objetos que ajudam seus usuários potenciais a terem vantagem completa do conhecimento da sua existência ou características.

Segundo DEXTER (2003), o prefixo *meta* vem do grego que significa junto, ao lado ou depois. Aplicando-o a um objeto significa que ele está em um estágio de desenvolvimento mais alto do que o objeto base. Ele é mais compreensivo e transcende a base. Em termos de dados, metadados transcende o dado de forma que ele vai além de um item de dado individual.

Os metadados não fornecem somente o formato e o nome dos itens de dado, ele fornece o contexto em que o elemento de dado existe. Fornecem também informações tais como o domínio dos possíveis valores, a relação do elemento de dado com outros, regras de negócio do dado e mesmo a origem do dado.

Para MORIARTY¹, citado por IKEMATU (2003), ao pesquisar sobre metadados, diversas interpretações são encontradas. Isso se deve ao fato de que essas diferentes contextualizações estão relacionadas ao estágio da organização, que os aplica dentro da hierarquia evolucionária da gestão do conhecimento, ou seja, organizações no nível mais baixo da hierarquia gerenciam dados brutos, enquanto organizações mais avançadas são capazes de administrar seus dados no nível de informações, conhecimento e sabedoria.

BONIFÁCIO (2003) define metadados, como uma referência de uma estrutura descritiva de informações sobre outro dado que é utilizado para auxiliar na definição, descrição e gerenciamento de recursos na *Web*.

Segundo POLLONI (2000, p. 27), metadados são “instrumentos para a busca de informação, mas no caso de documentos têm uma função adicional do ponto de vista arquivístico: atender aos requisitos de administração, como, por exemplo, a determinação do ciclo de vida e, portanto, o prazo de retenção dos documentos, base para decisões sobre localizações e meio de armazenamento, migração, etc.”.

Para SOUZA, CATARINO e SANTOS (1997, p. 93–94), metadados são circunscrições de dados estocados em um banco de dados, ou como geralmente é definido, “dados sobre dados armazenados a partir de um dicionário digital de dados”.

¹ MORIARTY, T. What is Metadata? **Database Programming and Design**., San Mateo, v. 10, n. 7, p. 57-59, July 1997.

De acordo com CAPPELI (2003), metadados é um dicionário de informações digitais que deve ser bem construído e conter informações sobre a origem do dado, regras de transformação, nomes, formatos de dados, entre outros, ou seja, deve conter muito mais do que descrições sobre colunas e tabelas, deve conter informação que adicione valor ao dado. Ainda segundo a autora, este dicionário de informação deve incluir: origem, fluxo do dado, formato do dado, nome, definições de negócios, regras de transformação, atualização e indicadores de qualidade de dados, requisitos de testes, responsabilidade sobre as informações, acesso e segurança.

Assim, conclui-se que os metadados são informações que têm como principal objetivo descrever um recurso informacional, auxiliando na sua identificação, caracterização, localização e utilização.

4.4.2 Função de Metadados

A tecnologia de metadados surgiu em função da necessidade das organizações em conhecerem os dados que elas mantêm.

De acordo com SOUZA, CATARINO e SANTOS (1997, p. 95) destacam-se as seguintes vantagens na utilização de recursos de metadados:

- a) estabelecimento de padrões de dados diante da heterogeneidade de informações contidas na rede;
- b) facilidade na definição da linguagem de consulta;
- c) facilidade e maior precisão na recuperação das informações desejadas;
- d) troca de informações entre aplicações e entre organizações;
- e) utilização das tecnologias na catalogação descritiva.

Para IKEMATU (2003), a principal finalidade dos metadados é documentar e organizar de forma estruturada os dados das organizações, com o objetivo de evitar duplicação de esforços e facilitar a manutenção dos dados, pois o autor ainda afirma que metadados é um dado associado a objetos que auxilia os usuários a terem vantagem completa do conhecimento da sua existência ou suas características.

Os metadados são responsáveis por fazerem administração dos dados, para garantirem a integridade e coesão dos mesmos, assim como também auxiliar na melhor utilização das funcionalidades de sistemas.

4.4.3 Padrões de Metadados

Os padrões de metadados têm como principal função fornecer definições comuns para a formação de uma rede automatizada de registros das propriedades e dos dados tornando-os consistentes, compatíveis e inteporáveis.

Para SOUZA, CATARINO e SANTOS (1997, p. 95), “é necessário e possível estabelecer padrões de metadados de forma que as organizações possam ser encorajadas a usá-los no sentido de contribuir para a documentação de seus dados”.

Segundo IKEMATU (2003), existem diferentes padrões de metadados para finalidades distintas. O autor afirma que os padrões de metadados de negócio são descrições de dados necessários pelos usuários de negócio, para atender o contexto e o significado dos dados. Os padrões de metadados se diferenciam conforme a finalidade do gerenciamento da informação desejada.

Alguns destes padrões são:

- a) *Directory Interchance Format* (DIF) – para criar entradas de diretórios que descrevem grupos de dados;
- b) *Government Information Locator Service* (GILS) – para informações governamentais;
- c) *Federal Data Geographic Committe* (FGDC) – para a descrição de dados geoespaciais;
- d) *Machine Readeble Card* (MARC) – para catalogação bibliográfica;
- e) *Dublin Core* (DC) – para dados sobre páginas da *Web*;
- f) *Consortium for the Interchange of Museum Information* (CIMI) – para informações sobre museus;
- g) *Meta Data Interchange Specification* (MDIS) - para troca de metadados entre ferramentas da tecnologia de informação;
- h) *Open Information Model* (OIM) – para facilitar o compartilhamento e reuso no desenvolvimento de aplicações e *data warehouse*;
- i) *Common Warehouse Meta Model* (CWM) – para troca de informações entre esquemas de banco de dados e *data warehouse*.

A seguir será apresentado o padrão *Dublin Core*, que pode ser definido como sendo o conjunto de elementos de metadados planejado para facilitar a descrição de documentos eletrônicos na *Web*.

4.4.3.1 Padrão *Dublin Core*

Para DESIO (2003), o *Dublin Core Metadata Element Set* ou simplesmente *Dublin Core*, é um padrão internacional para descrição de recursos de informação, considerado parte importante da infra-estrutura da *Internet*. Foi desenvolvido a partir de 1994, por um grupo de bibliotecários e especialistas de conteúdo, encabeçado por Stuart Weibel, pesquisador da *Online Computer Library Center* (OCLC). É chamado *Dublin Core* por ter se originado em um *workshop* na cidade de Dublin, Ohio, Estados Unidos. Todas as ações e esforços são liderados pelo *Dublin Core Metadata Initiative* (DCMI) que é o responsável pelo desenvolvimento, padronização e promoção do conjunto de elementos metadados *Dublin Core*.

Embora o *Dublin Core* possa ser usado para descrever materiais em formatos tradicionais, o seu impulso maior tem sido para descrever recursos eletrônicos disponíveis na *Internet*. Consiste em um conjunto de 15 (quinze) elementos de metadados, equivalentes a uma ficha catalográfica, os quais podem ser considerados como o mais baixo denominador comum para descrição de recursos de informação (WEIBEL, 2003).

As principais características do padrão *Dublin Core* são:

- a) simplicidade na descrição dos recursos: pode ser usado por não-catalogadores, autores ou *websiters* sem conhecimento prévio de todas as regras de catalogação;
- b) interoperabilidade semântica: promove o entendimento comum dos descritores, ajudando a unificar padrões de conteúdo, aumentando a possibilidade de interoperabilidade semântica entre as disciplinas;
- c) consenso internacional: reconhecimento da cobertura internacional do escopo do recurso;
- d) extensibilidade: constitui uma alternativa aos modelos de descrição mais elaborados e caros. Possui flexibilidade e extensibilidade para codificar semânticas mais elaboradas em padrões mais sofisticados.

Atualmente, são inúmeros os projetos em desenvolvimento nos diversos continentes (América do Norte, Europa, Ásia e Austrália), que utilizam o padrão *Dublin Core*. Entre as iniciativas pioneiras de desenvolvimento de ferramentas para descrição de recursos eletrônicos, tomando por base o *Dublin Core*, destacam-se os projetos:

- a) *The Nordic Metadata Project*: primeiro projeto internacional a optar pelo *Dublin Core* no desenvolvimento de ferramentas para geração, coleta e indexação de metadados, envolvendo países escandinavos: Finlândia, Noruega, Dinamarca, Suécia e Islândia;
- b) *The Donor Project*: desenvolvido pela Biblioteca Nacional da Holanda, adota o padrão *Dublin Core* na gestão de metadados e possui ferramenta para gerar metadados;
- c) *Foundations Project Minnesota's Gateway to Environmental Information*: com a colaboração de agências ambientais do governo do Estado de Minnesota, Estados Unidos, utiliza o padrão para geração de seus próprios metadados.

Desde 1996, o *Dublin*, vem se evidenciando como a solução mais viável para descrição de recursos eletrônicos na *Internet*. Projetos de construção de ferramentas baseadas no *Dublin Core* têm aumentado significativamente, assim como a quantidade de documentos disponíveis na *Internet* que contêm suas próprias descrições de *Dublin Core*.

4.4.4 Linguagens de marcação para descrição de metadados

4.4.4.1 Hypertext Markup Language (HTML)

A linguagem básica utilizada para as publicações na *Web* é o *Hypertext Markup Language* (HTML), uma derivação de *Standard Generalized Markup Language* (SGML). O HTML foi desenvolvido para ser simples, tanto para os autores dos documentos de marcação de hipertextos como para os desenvolvedores de *browsers*, os quais são os interpretadores dos códigos HTML. Atualmente, o HTML se encontra na versão 4.0, e conta com quase 100 *tags*² que definem a estrutura dos textos de marcação. A linguagem HTML possibilita a construção de documentos por meio da descrição de seus três componentes básicos:

- a) conteúdo dos dados: trata das palavras propriamente ditas do documento;
- b) estrutura dos dados: define os elementos que compõem o texto;
- c) apresentação dos dados: define a forma como os dados serão apresentados, tipo de letra, tabulação, texto, voz, cor, etc.

Contudo, ocorrem problemas na codificação de documentos HTML, pois todos os componentes básicos do documento são codificados em conjunto com a digitação do texto, não havendo nenhuma forma estruturada para representar estes elementos em separado.

O escopo da *Internet* estendeu-se de tal forma que a simplicidade da HTML está se tornando seu ponto mais fraco no atendimento a demanda de novas áreas como: comércio eletrônico, saúde, educação e outros serviços *on-line*. Com isso, demonstrou-se necessária a criação de um padrão formal, mais robusto, definido e flexível que pudesse atender às características de cada uma destas aplicações e serviços que são disponibilizados na *Web*.

Para ROCHA (2003), tecnologias como sistemas gerenciadores de banco de dados e linguagens de programação, são usadas em conjunto com HTML, pela grande maioria dos desenvolvedores de sistemas e páginas na *Internet* visando uma melhor estruturação, armazenamento e recuperação das informações. Sem o auxílio destas tecnologias o uso e a manutenção das informações armazenadas na *Internet*

² *Tags* são marcas utilizadas pelas linguagens de marcação, que informam o limite e a função do dado/informação no documento/recurso.

seriam prejudicados, tendo em vista que o HTML somente oferece recursos para armazenamento de textos, sem quaisquer mecanismos para execução de programas e estruturação dos dados.

Visando potencializar a interatividade entre os usuários e aumentar a integração dos sistemas na *Internet* à banco de dados, foram incorporadas algumas soluções como:

- a) *Common Gateway Interface* (CGI): é um padrão para o interfaceamento das aplicações externas com os servidores de informação, tais como HTTP ou Servidores *Web*. A forma de recuperação de informações de um documento HTML é estática, o que significa que existe um estado constante: um arquivo texto que não muda. Um programa em CGI, por outro lado, é executado de forma *on-line*, o que permite recuperar informações dinamicamente;
 - b) Java: linguagem de programação que independe da plataforma onde a aplicação resultante estará sendo executada. Após um programa escrito em Java a ser compilado ele pode ser executado em qualquer plataforma *WIN32*, *Solais*, *Maços*, etc, desde que possua um interpretador para a linguagem Java;
 - c) *Active Server Pages* (ASP): ambiente para programação por *scripts* no servidor, que pode ser usado para criação de páginas dinâmicas, interativas e de alta performance. Como as páginas ASP, os *scripts* rodam no servidor e não no cliente. É o próprio servidor que transforma os *scripts* em HTML padrão, fazendo com que qualquer *browser* do mercado seja capaz de acessar um *site* que usa ASP.
 - d) *Practical Extraction and Report Language* (PERL): linguagem de programação criada por *Larry Wall*, versátil para o desenvolvimento de programas que necessitem de tratamento de *strings* e a mais adequada para criação de *scripts* CGI nos servidores *Internet*;
 - e) *Personal Home Page*: desenvolvida em 1994 por Rasmus Lerdorf, é uma linguagem para manipulação de *scripts* desenvolvida a partir de C, Java e PERL, visando facilitar a criação de páginas dinâmicas na *Web*.
-

4.4.4.2 Extensible Mark-up Language (XML)

Implementadores *Web* estão diante do aumento do número de produtos e fontes de dados usados no seu ambiente. Implementar portais de informações, por exemplo, implica no uso de um grande número de fonte de dados e seus metadados. Uma das tendências em ambientes cada vez mais distribuídos é alavancar meios de armazenamento. A única solução viável, neste ambiente, é fornecer gerenciamento de metadados distribuídos facilitando o fluxo de informações entre sistemas diferentes. Neste ambiente, padrões de intercâmbio de metadados são a melhor escolha.

Para FINKELSTEIN (2003) “o XML é um subconjunto do SGML, que tenta definir um padrão universal para troca eletrônica de dados. O XML mantém aspectos arquiteturais importantes de separação de contexto removendo características não essenciais de SGML.” O XML é um padrão de tecnologia aberta do *World Wide Web Consortium* (W3C), o grupo responsável pela manutenção e evolução do HTML e outros padrões relacionados com a Web.

O XML oferece um mecanismo baseado em texto para representar a estrutura do dado original e interpretar este dado de forma não ambígua. Ele usa uma abordagem semelhante ao HTML para manusear a apresentação dos conteúdos. Entretanto, o XML descreve a característica do dado que ele especifica, além de ser mais flexível e amplo que o HTML no tratamento de metadados associados com o conteúdo da informação.

Os documentos XML são bens de informação, tal como bases de dados e modelos lógicos. Da mesma forma, eles exigem um gerenciamento apropriado de metadados. É importante entender os documentos XML que são usados na organização e verificar o impacto das mudanças nestes documentos para organização.

Segundo BOULET (2003), o XML permite melhor gerenciamento e reuso de documentos estruturados. Isto é feito por meio da separação do conteúdo da *Internet* em um repositório de dados da lógica de negócio (regras que governam o conteúdo disponibilizado) e apresentação (os modelos de *layout* que determinam a aparência gráfica do *site*). Por meio desta separação, o XML facilita o uso do conteúdo de várias formas e com várias apresentações diferentes. A separação do

conteúdo, apresentação e lógica facilita o processo de disponibilizar o conteúdo. Páginas *Internet* podem agora ser construídas baseadas na ação do visitante. Quando um pedido é feito pelo usuário, os dados que o usuário pode ver são obtidos do repositório baseado nas regras de negócio definidas.

Segundo CUTLIP (2003), bancos de dados relacionais com extensões XML podem fornecer um estrutura básica robusta e flexível para metadados. As extensões XML fornecem suporte de infra-estrutura para gerenciamento de dados e metadados através de XML. A finalidade básica do XML é manipular e mapear as estruturas hierárquicas do XML para colunas de tabelas relacionais.

Há condições também de uso do *XSL Transformation* (XSLT), processador XSLT pode ser usado simultaneamente com o XML para traduzir mensagens entre bases de dados heterogêneas.

Segundo LEWIS (2000, p. 34), extensões XML fornecem a facilidade de mapeamento SQL para criar documentos XML de duas ou mais tabelas incluindo uso de chaves primária e estrangeira. Produtos como *Microsoft SQL Server 2000* e novas soluções de sistemas gerenciadores de banco de dados baseados em XML como o *Excelon* e o Tamino da *Software AG* seguem esta tendência.

5 ANÁLISE DOS DADOS

Nesta seção serão apresentados os resultados do levantamento de dados feito em campo, após o levantamento da literatura pertinente.

Verificou-se que 36 (92,3%) das IES da rede particular de Curitiba possuem *sites* ou páginas na *Internet*, e 3 (7,7%) instituições não disponibiliza qualquer página, conforme pode-se verificar no quadro 1 e gráfico 1.

Observa-se que várias instituições compartilham o mesmo *site*, influenciando nesses números. São elas: Centro de Educação Tecnológica Martinus (FELC) e Faculdade Evangélica Luterana de Curitiba (FELC), Centro de Educação Tecnológica OPET (CET) e Faculdade Organização Paranaense de Ensino Técnico (FAO), Faculdade de Ciências e Tecnologia do Paraná (FACET-PR) e Faculdade de Ciências Aplicadas do Paraná, Faculdade Padre João Bagozzi - Faculdade Bagozzi e Instituto Superior de Educação Padre João Bagozzi, Faculdade Paranaense de Administração (FPA) e Faculdades Spei (FACSPEI) e Faculdades Integradas Santa Cruz de Curitiba – FARESC e Instituto Sul Brasil de Ensino Superior.

Exibindo o código fonte da *home page* de cada um dos *sites* pesquisados, detectou-se que 30 (83,3%) deles utiliza a tecnologia de metadados, conforme mostram o quadro 1 e o gráfico 2.

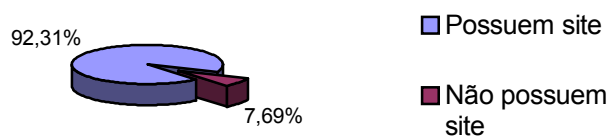
Quando se pesquisou sobre o padrão de metadados, verificou-se que todos os que possuem essa tecnologia (30 *sites*) utilizam o simples padrão HTML, conforme verifica-se no quadro 1, complementado pelo quadro 2 que apresenta quais os metadados utilizados.

QUADRO 1 – EXISTÊNCIA DE SITES E METADADOS DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA NA INTERNET - 2003

INSTITUIÇÃO	SITE	METADADOS	
		POSSUI	PADRÃO
Centro de Educação Tecnológica Camões – CETESC	-	-	-
Centro de Educação Tecnológica Internacional – CENINTER	x	x	Html
Centro de Educação Tecnológica Martinus – FELC	x	x	Html
Centro de Educação Tecnológica OPET – CET	x	x	Html
Centro Universitário Campos de Andrade – UNIANDRADE	x	x	Html
Centro Universitário Positivo- UNICENP	x	-	-
Escola Superior de Estudos Empresariais e Informática – ESEEI	x	-	-
Escola Superior de Gestão Comercial e Marketing – ESIC	x	x	Html
Faculdade Anchieta – FAESP	x	-	-
Faculdade Camões	-	-	-
Faculdade de Administração, Ciências, Educação e Letras – FACEL	x	x	Html
Faculdade de Ciências e Tecnologia do Paraná - FACET-PR	x	x	Html
Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET	x	x	Html
Faculdade de Filosofia São Boaventura	x	x	Html
Faculdade Dom Bosco – FDB	x	x	Html
Faculdade Doutor Leocádio José Correia – FALEC	x	x	Html
Faculdade Evangélica do Paraná – FEPAR	x	x	Html
Faculdade Evangélica Luterana de Curitiba – FELC	x	x	Html
Faculdade Hoyer de Comunicação Social – FHCS	x	x	Html
Faculdade Internacional de Curitiba – FACINTER	x	x	Html
Faculdade Organização Paranaense de Ensino Técnico – FAO	x	x	Html
Faculdade Padre João Bagozzi - Faculdade Bagozzi	x	x	Html
Faculdade Paranaense de Administração – FPA	x	x	Html
Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	x	-	-
Faculdade Teológica Batista do Paraná – FTBP	x	x	Html
Faculdades Bom Jesus – FBJ	x	x	Html
Faculdades Integradas "Espirita" – FIES	x	x	Html
Faculdades Integradas Curitiba – FIC	x	-	-
Faculdades Integradas do Brasil – UNIBRASIL	x	x	Html
Faculdades Integradas Santa Cruz de Curitiba – FARESC	x	x	Html
Faculdades SPEI – FACSPEI	x	x	Html
Instituto de Ciências Sociais do Paraná- ICSP	x	x	Html
Instituto de Ensino e Cultura do Paraná – IECF	-	-	-
Instituto Sul Brasil de Ensino Superior	x	x	Html
Instituto Superior de Educação Nossa Senhora de Sion - ISE – Sion	x	-	-
Instituto Superior de Educação Padre João Bagozzi	x	x	Html
Pontifícia Universidade Católica do Paraná – PUC-PR	x	x	Html
Unidade de Ensino Superior Expoente – UNIEXP	x	x	Html
Universidade Tuiuti do Paraná – UTP	x	x	Html

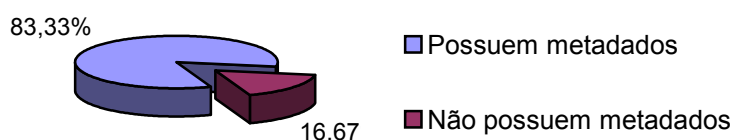
FONTE: Elaboração do autor

GRÁFICO 1 – INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM SITES - 2003



FONTE: Elaboração do autor

GRÁFICO 2 –SITES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM METADADOS - 2003



FONTE: Elaboração do autor

Quanto à linguagem de marcação utilizada pelas IES da rede particular de Curitiba, denotou-se que 4 (11,1%) usam somente a linguagem HTML, 26 (72,2%) utilizam as linguagens HTML e Javascript, 3 (5,5%) fazem uso das linguagens HTML, Javascript e ASP, 1 (2,8%) usa as linguagens HTML e ASP, 1 (2,8%) utiliza as linguagens HTML, Javascript e PHP e 1 (2,8%) faz uso somente da linguagem Javascript, conforme pode-se observar no quadro 3.

QUADRO 2 - METADADOS HTML UTILIZADOS PELAS *HOME PAGES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA - 2003

continua

INSTITUIÇÃO	PADRÃO HTML
Centro de Educação Tecnológica Camões – CETESC	-
Centro de Educação Tecnológica Internacional – CENINTER	Título, Palavras - chave
Centro de Educação Tecnológica Martinus – FELC	Título, Descrição, Palavras - chave, Autor
Centro de Educação Tecnológica OPET – CET	Título
Centro Universitário Campos de Andrade – UNIANDRADE	Título
Centro Universitário Positivo- UNICENP	-
Escola Superior de Estudos Empresariais e Informática – ESEEI	-
Escola Superior de Gestão Comercial e Marketing – ESIC	Título, Descrição, Palavras - chave, Autor
Faculdade Anchieta – FAESP	-
Faculdade Camões	-
Faculdade de Administração, Ciências, Educação e Letras – FACEL	Título
Faculdade de Ciências e Tecnologia do Paraná - FACET-PR	Título, Descrição, Palavras - chave, Autor
Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET	Título, Descrição, Palavras - chave, Autor
Faculdade de Filosofia São Boaventura	Título
Faculdade Dom Bosco – FDB	Título, Identificador, Descrição, Idioma, Palavras-chave, Autor, Direitos
Faculdade Doutor Leocádio José Correia – FALEC	Título
Faculdade Evangélica do Paraná – FEPAR	Título
Faculdade Evangélica Luterana de Curitiba – FELC	Título, Descrição, Palavras - chave, Autor
Faculdade Hoyler de Comunicação Social – FHCS	Título
Faculdade Internacional de Curitiba – FACINTER	Título
Faculdade Organização Paranaense de Ensino Técnico – FAO	Título
Faculdade Padre João Bagozzi - Faculdade Bagozzi	Título
Faculdade Paranaense de Administração – FPA	Título
Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	-
Faculdade Teológica Batista do Paraná – FTBP	Título
Faculdades Bom Jesus – FBJ	Título, Descrição, Palavras - chave, Criador, Idioma, Autor
Faculdades Integradas "Espirita" – FIES	Título
Faculdades Integradas Curitiba – FIC	-
Faculdades Integradas do Brasil – UNIBRASIL	Título, Palavras - chave
Faculdades Integradas Santa Cruz de Curitiba – FARESC	Título, Palavras - chave

QUADRO 2 - METADADOS HTML UTILIZADOS PELAS *HOME PAGES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA - 2003

conclusão

INSTITUIÇÃO	PADRÃO HTML
Faculdades SPEI – FACSPEI	Título
Instituto de Ciências Sociais do Paraná- ICSP	Título
Instituto de Ensino e Cultura do Paraná – IECP	-
Instituto Sul Brasil de Ensino Superior	Título, Palavras - chave
Instituto Superior de Educação Nossa Senhora de Sion - ISE – Sion	-
Instituto Superior de Educação Padre João Bagozzi	Título
Pontifícia Universidade Católica do Paraná – PUC-PR	Título
Unidade de Ensino Superior Expoente – UNIEXP	Título
Universidade Tuiuti do Paraná – UTP	Título, descrição, Palavras-chave, Autor

FONTE: Elaboração do autor

QUADRO 3 – IDENTIFICAÇÃO DAS LINGUAGENS DE MARCAÇÃO UTILIZADAS NOS SITE DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA – 2003

INSTITUIÇÃO	LINGUAGEM
Centro de Educação Tecnológica Camões – CETESC	-
Centro de Educação Tecnológica Internacional – CENINTER	html, javascript
Centro de Educação Tecnológica Martinus – FELC	html, javascript, asp
Centro de Educação Tecnológica OPET – CET	html, javascript
Centro Universitário Campos de Andrade – UNIANDRADE	html, asp
Centro Universitário Positivo- UNICENP	html, javascript, asp
Escola Superior de Estudos Empresariais e Informática – ESEEI	html, javascript
Escola Superior de Gestão Comercial e Marketing – ESIC	html, javascript
Faculdade Anchieta – FAESP	html, javascript
Faculdade Camões	-
Faculdade de Administração, Ciências, Educação e Letras – FACEL	html, javascript
Faculdade de Ciências e Tecnologia do Paraná - FACET-PR	html, javascript
Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET	html, javascript
Faculdade de Filosofia São Boaventura	html
Faculdade Dom Bosco – FDB	html, javascript
Faculdade Doutor Leocádio José Correia – FALEC	html
Faculdade Evangélica do Paraná – FEPAR	html, javascript
Faculdade Evangélica Luterana de Curitiba – FELC	html, javascript, asp
Faculdade Hoyler de Comunicação Social – FHCS	html, javascript
Faculdade Internacional de Curitiba – FACINTER	html, javascript
Faculdade Organização Paranaense de Ensino Técnico – FAO	html, javascript
Faculdade Padre João Bagozzi - Faculdade Bagozzi	html, javascript
Faculdade Paranaense de Administração – FPA	html, javascript
Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	html, javascript
Faculdade Teológica Batista do Paraná – FTBP	html
Faculdades Bom Jesus – FBJ	html
Faculdades Integradas "Espirita" – FIES	html, javascript
Faculdades Integradas Curitiba – FIC	javascript
Faculdades Integradas do Brasil – UNIBRASIL	html, javascript
Faculdades Integradas Santa Cruz de Curitiba – FARESC	html, javascript
Faculdades SPEI – FACSPEI	html, javascript
Instituto de Ciências Sociais do Paraná- ICSP	html, javascript, php
Instituto de Ensino e Cultura do Paraná – IECP	-
Instituto Sul Brasil de Ensino Superior	html, javascript
Instituto Superior de Educação Nossa Senhora de Sion - ISE – Sion	html, javascript
Instituto Superior de Educação Padre João Bagozzi	html, javascript
Pontifícia Universidade Católica do Paraná – PUCPR	html, javascript
Unidade de Ensino Superior Expoente – UNIEXP	html, javascript
Universidade Tuiuti do Paraná – UTP	html, javascript

FONTE: Elaboração do autor

Observou-se que o metadado “Título” é o mais utilizado nos *sites* das IES da rede particular de Curitiba, no entanto como este metadado é visto apenas como uma forma de identificação do *site* e não tem o objetivo de toar como critério de busca, optou-se por utilizar o metadado “palavras-chave” que foi o segundo mais utilizado, conforme mostra o quadro 2. Foram identificados os termos de seus conteúdos que mais se repetiram, destacando-se 6 (seis) entre eles: Faculdade, Ensino, Vestibular, Curso Superior, Educação e Superior, conforme se observa no quadro 4.

Realizou-se, também, um levantamento dos buscadores mais acessados no Brasil, identificando-se o *Google*, o Cadê e o Altavista, com a finalidade de classificar os *sites* das IES pesquisadas, segundo a colocação no resultado da busca na *Internet*.

Foram então realizadas buscas nas ferramentas identificadas acima, utilizando-se dos 6 termos destacados no metadado “palavra-chave” com o objetivo de analisar a posição ocupada por cada IES, para assim saber se o metadado está fazendo alguma diferença no mercado virtual.

Como resultado destas buscas, realizadas na opção Busca Avançada das ferramentas, *Google*, Cadê e Altavista, constatou-se que, conforme mostra o quadro 5, que pelo termo Faculdade não se recuperou nenhuma das IES nos primeiros duzentos itens, sendo que esperava-se que todos aparecessem por utilizar o recurso metadado, e isso ocorreu também com os termos Curso Superior e Educação. Já com os termos Ensino, Superior e Vestibular, houve a identificação de 5 IES.

Alertando que destes, nem sempre, aquelas que apareceram possuem o termo correspondente, conforme pode se comparar com o quadro 4. Observou-se esta ocorrência na utilização do termo Ensino onde verifica-se que, a Universidade Tuiuti do Paraná (UTP) que possui este termo apareceu em 164.º, já as Faculdades Integradas Santa Cruz de Curitiba (FARESC) e o Instituto Sul Brasil de Ensino Superior que não possuem este termo também apareceram em 83.º.

QUADRO 4 - INCIDÊNCIA DOS TERMOS NO METADADO “PALAVRA – CHAVE” DOS SITES DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA - 2003

INSTITUIÇÃO	PALAVRAS - CHAVE					
	FACULDADE	ENSINO	VESTIBULAR	CURSO SUPERIOR	EDUCAÇÃO	SUPERIOR
Centro de Educação Tecnológica Internacional – CENINTER	X	X	X	-	X	X
Centro de Educação Tecnológica Martinus – FELC	X	-	-	-	-	-
Escola Superior de Gestão Comercial e Marketing – ESIC	X	X	X	-	-	X
Faculdade de Ciências e Tecnologia do Paraná - FACET-PR	X	X	-	-	-	-
Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET	X	X	-	-	-	-
Faculdade Dom Bosco – FDB	X	-	X	-	X	-
Faculdade Evangélica Luterana de Curitiba – FELC	X	-	-	-	-	-
Faculdades Bom Jesus – FBJ	X	X	-	-	-	X
Faculdades Integradas do Brasil – UNIBRASIL	X	-	-	-	-	-
Faculdades Integradas Santa Cruz de Curitiba – FARESC	X	-	-	X	-	-
Instituto Sul Brasil de Ensino Superior	X	-	-	X	-	-
Universidade Tuiuti do Paraná – UTP	X	X	X	X	X	-

FONTE: Elaboração do autor

QUADRO 5 – POSIÇÃO DOS *SITES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” - 2004

TERMOS	INSTITUIÇÕES	BUSCADOR/COLOCAÇÃO		
		GOOGLE	CADÊ	ALTAVISTA
Curso Superior	-	-	-	-
Educação	-	-	-	-
Ensino	Faculdades Integradas Santa Cruz de Curitiba – FARESC	-	-	83.º
	Instituto Sul Brasil de Ensino Superior	-	-	83.º
	Universidade Tuiuti do Paraná – UTP	-	-	164.º
Faculdade	-	-	-	-
Superior	Universidade Tuiuti do Paraná – UTP	-	147.º	-
	Escola Superior de Gestão Comercial e Marketing - ESIC	-	-	70.º
Vestibular	Faculdades Integradas do Brasil - UNIBRASIL	134.º	-	-

FONTE: Elaborado pelo autor

QUADRO 6 - POSIÇÃO DOS *SITES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE NÃO POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” - 2004

TERMOS	INSTITUIÇÕES	BUSCADOR/COLOCAÇÃO		
		GOOGLE	CADÊ	ALTAVISTA
Curso Superior	-	-	-	-
Educação	-	-	-	-
Ensino	Escola Superior de Estudos Empresariais e Informática - ESSEI	-	-	69.º
Faculdade	Escola Superior de Estudos Empresariais e Informática - ESSEI	-	16.º	-
	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	-	60.º	173.º
Superior	-	-	-	-
Vestibular	-	-	-	-

FONTE: Elaborado pelo autor

Realizou-se também uma busca com as mesmos 6 termos destacados do quadro 4, com o intuito de identificar a posição das IES que não possuem metadados, frente ao resultado da busca conforme observa-se no quadro 6. Observou-se que apenas duas das IES que não possuem metadados apareceram. No entanto, estabelecendo uma comparação entre os quadros 5 e 6, em relação aos termos Faculdade e Ensino, verifica-se que à posição ocupada no *ranking* das buscas, as IES que não possuem metadados ocuparam às primeiras colocações frente aos resultados das pesquisas.

Como o índice de recuperação dos *sítes* das IES denotadas foi pequeno, optou-se por uma nova busca aplicando os mesmos termos até então utilizados, unido ao operador booleano “E” , pois segundo GUINCHAT e MENOU (1994, p. 306) este operador “ permite relacionar dois descritores que devem ser encontrados na indexação de um mesmo documento para que este documento seja considerado pertinente” .

A busca foi realizada com os 6 termos e unidos à palavra-chave Curitiba como fator de seleção, já que todas as IES pertencem a esta cidade, para verificar se o índice de recuperação aumentava.

Primeiramente foi realizada as buscas para identificar as IES que possuem metadados. Foram utilizados os 6 termos e a palavra Curitiba, observou-se que em comparação com as buscas explicitadas no quadro 5 que esse índice aumentou significativamente, conforme verifica-se no quadro 7.

Ainda estabelecendo esta comparação entre os quadros 5 e 7 observa-se que no termo Faculdade na primeira busca não se obteve nenhum resultado ao contrário da segunda, em que foi acrescentado a palavra-chave Curitiba, observou-se que houve um aumento recuperação para 9 *sítes*, assim ocorreu também em comparação as outras buscas. Lembrando que na primeira busca houve apenas 5 recuperações, já na segunda o índice de recuperação aumentou 37 incidências .

QUADRO 7 – POSIÇÃO DOS *SITES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” - 2004

continua

TERMOS	INSTITUIÇÕES	BUSCADOR/COLOCAÇÃO		
		GOOGLE	CADÊ	ALTAVISTA
Curso Superior + Curitiba	Centro de Educação Tecnológica Martinus – FELC	28.º	-	-
	Faculdade Evangélica Luterana de Curitiba – FELC	28.º	-	-
	Universidade Tuiuti do Paraná – UTP	-	17.º	-
	Faculdades Integradas Santa Cruz de Curitiba – FARESC	-	49.º	-
	Instituto Sul Brasil de Ensino Superior	-	49.º	-
	Faculdade de Ciências e Tecnologia do Paraná - FACET-PR	-	198.º	-
	Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET	-	198.º	-
	Faculdades Integradas do Brasil – UNIBRASIL	-	-	138.º
Educação + Curitiba	Faculdades Integradas Santa Cruz de Curitiba – FARESC	144.º	-	30.º
	Instituto Sul Brasil de Ensino Superior	144.º	-	30.º
	Universidade Tuiuti do Paraná – UTP	84.º	-	-
	Centro de Educação Tecnológica Martinus – FELC	-	156.º	-
	Faculdade Evangélica Luterana de Curitiba – FELC	-	156.º	-
Ensino + Curitiba	Faculdade Dom Bosco – FDB	5.º	2.º	152.º
	Faculdades Integradas Santa Cruz de Curitiba – FARESC	7.º	4.º	15.º
	Instituto Sul Brasil de Ensino Superior	7.º	4.º	15.º
	Universidade Tuiuti do Paraná – UTP	51.º	-	167.º

QUADRO 7 – POSIÇÃO DOS *SITES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” - 2004

conclusão

TERMOS	INSTITUIÇÕES	BUSCADOR/COLOCAÇÃO		
		GOOGLE	CADÊ	ALTAVISTA
Faculdade + Curitiba	Centro de Educação Tecnológica Martinus – FELC	118.º	24.º	2.º
	Faculdade Evangélica Luterana de Curitiba – FELC	118.º	24.º	2.º
	Faculdade Dom Bosco – FDB	12.º	3.º	-
	Faculdades Integradas Santa Cruz de Curitiba – FARESC	23.º	4.º	-
	Instituto Sul Brasil de Ensino Superior	23.º	4.º	-
	Escola Superior de Gestão Comercial e Marketing – ESIC	-	115.º	56.º
	Faculdades Integradas do Brasil – UNIBRASIL	-	-	3.º
	Faculdade de Ciências e Tecnologia do Paraná - FACET-PR	-	-	79.º
	Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET	-	-	79.º
Superior + Curitiba	Escola Superior de Gestão Comercial e Marketing – ESIC	-	-	9.º
	Faculdade Dom Bosco – FDB	-	-	31.º
	Centro de Educação Tecnológica Martinus – FELC	-	-	97.º
	Faculdade Evangélica Luterana de Curitiba – FELC	-	-	97.º
	Faculdades Integradas Santa Cruz de Curitiba – FARESC	-	180.º	-
	Instituto Sul Brasil de Ensino Superior	-	180.º	-
Vestibular + Curitiba	Faculdades Integradas Santa Cruz de Curitiba – FARESC	38.º	45.º	22.º
	Instituto Sul Brasil de Ensino Superior	38.º	45.º	22.º
	Faculdades Integradas do Brasil – UNIBRASIL	24.º	-	110.º
	Faculdade Dom Bosco – FDB	-	15.º	6.º
	Universidade Tuiuti do Paraná – UTP	12.º	-	-

FONTE: Elaboração do autor

Por fim, realizou-se a mesma pesquisa, utilizando os 6 termos unidos à palavra Curitiba, mas agora, com o objetivo de identificar dentre aqueles *sítes* que não possuem o padrão metadados, com a finalidade de verificar se o índice de amostragem aumentava em comparação com o quadro 6. Observou-se que este índice cresceu de 4 para 14 recuperações, conforme verifica-se no quadro 8.

Concluiu-se que, após os levantamentos e tabulações dos dados, atualmente as ferramentas de buscas provavelmente utilizando o metadado como critério de busca e recuperação.

QUADRO 8 - POSIÇÃO DOS *SITES* DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA QUE NÃO POSSUEM METADADOS NOS RESULTADOS DAS BUSCAS UTILIZANDO OS TERMOS DO METADADO “PALAVRA-CHAVE” - 2004

TERMOS	INSTITUIÇÕES	BUSCADOR/COLOCAÇÃO		
		GOOGLE	CADÊ	ALTAVISTA
Curso Superior + Curitiba	Faculdades Integradas Curitiba – FIC	15.º	72.º	-
	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	72.º	-	-
Educação + Curitiba	Faculdades Integradas Curitiba – FIC	55.º	33.º	5.º
	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	-	-	60.º
Ensino + Curitiba	Faculdades Integradas Curitiba – FIC	2.º	1.º	1.º
	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	64.º	-	45.º
Faculdade + Curitiba	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	4.º	1.º	5.º
	Faculdades Integradas Curitiba – FIC	1.º	1.º	-
	Centro Universitário Positivo- UNICENP	-	116.º	-
Superior + Curitiba	Faculdades Integradas Curitiba – FIC	-	5.º	1º
	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	51.º	-	71.º
Vestibular + Curitiba	Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras	84.º	118.º	100.º
	Faculdades Integradas Curitiba – FIC	4.º	1.º	106.º
	Centro Universitário Positivo- UNICENP	160.º	-	-

FONTE: Elaboração do autor

6 CONSIDERAÇÕES FINAIS

O conhecimento e a informação se tornaram um bem precioso para a organização. Sendo assim, é necessário gerenciá-los bem para se ter um efetivo controle sobre estes patrimônios peculiares. O valor deste gerenciamento aumenta considerando que as informações estão cada vez mais integradas e passam a dar apoio a tomadas de decisão para que a organização obtenha vantagens competitivas também no mercado virtual.

Este trabalho procurou levantar os principais esforços existentes na área de tratamento de metadados. Embora este assunto seja tratado há muito tempo, encontra-se pouca literatura na língua portuguesa, em comparação com outras áreas da tecnologia da informação, esta dificuldade em encontrar material sobre metadados se transformou em um bom desafio para o desenvolvimento desta pesquisa.

Pode-se dizer que hoje a *Internet* sofre do mal da desordem informacional, com isso tem-se como consequência a dificuldade do gerenciamento estratégico da informação. Para isto, a literatura nos traz muitos documentos dando ênfase no recurso metadados, que é visto como uma forma de sanar alguns problemas da *Internet*.

Por meio desta pesquisa pode-se perceber que as IES da rede particular de Curitiba utilizam apenas o simples metadado HTML, e por mais que se utilize deste recurso, detectou-se que as ferramentas de busca não adotam estruturas na recuperação das informações no metadado dos *sites*, fazendo com que a representatividade das IES no mercado virtual fique prejudicada.

Com os resultados desta pesquisa exploratória percebeu-se que se fazem necessárias algumas recomendações associadas à literatura pertinente, onde destaca-se as seguintes:

- a) o termo metadado deve ser mais difundido e estudado no Brasil, afim de que haja uma maior conscientização dos benefícios que esta ferramenta pode trazer para o mercado virtual;
- b) equipes multidisciplinares precisam aliar seus conhecimentos, a fim de estimular a disseminação da informação de qualidade e da utilização de tecnologias adequadas;

- c) as ferramentas de busca devem desenvolver sistemas inteligentes associadas ao recurso metadados, por meio de “robôs” que façam esta busca nos metadados dos *sites*, ou que exista um campo específico no buscador para a pesquisa direta nos metadados.

Com este trabalho percebeu-se que o recurso metadado deve ser associado a outras tecnologias, como o XML e sistemas inteligentes de busca, pois estes poderiam auxiliar na recuperação mais precisa das informações e na vantagem competitiva no mercado virtual.

REFERÊNCIAS

- ANSOFF, H. I. **A nova estratégia empresarial**. São Paulo: Atlas, 1990.
- BACCHIN, T. R. **Os diferentes tipos de sistemas de busca**. Disponível em <<http://www.widebiz.com.br/gente/tbacchin/metatags.html>>. Acesso em: 12 dez. 2003.
- BEUREN, I. M. **Gerenciamento da informação**: um recurso estratégico no processo de gestão. 2. ed. São Paulo: Atlas, 2000.
- BONIFÁCIO, A. S. **Metadados semânticos para buscas em bibliotecas digitais**. Disponível em: <<http://www.uel.br/pessoal/ailton/Trabalhos/SemanaAcad-Ailton.html>>. Acesso em: 09 jul. 2003.
- BOULET, L. **XML to the rescue**: improving the efficiency of web content management. Disponível em: <http://www.dmreview.com/editorial/dmdirect_article_archive.cfm>. Acesso em: 10 out. 2003.
- CAPPELI, C. **Metadados**. Disponível em: <<http://genesis.nce.ufrj/dataware/Metados.htm>>. Acesso em: 18 jul. 2003.
- CASTELLS, M. **A sociedade em rede**. 3. ed. São Paulo: Paz e Terra, 2000. (A Era da Informação: economia, sociedade e cultura).
- CÉNDON, B. V. Ferramentas de busca na Web. **Ciência da Informação**, Brasília, v.30, n.1, p.39-49, jan./abr. 2001.
- CHOO, C. W. **Information management for the intelligent organization**: the art of scanning the environment. 2. ed. New York: Asis, 1998.
- CIANCONI, R. **Gestão da informação na sociedade do conhecimento**. 2. ed. Brasília: Brasil, 2001.
- CUTLIP, R. **XML extensers**: more than just data. Disponível em: <http://www.dmreview.com/editorial/dmdirect_article_archive.cfm?EdID=2309>. Acesso em: 10 out. 2003.
- DESIO, V. **Information metadata**. Disponível em: <<http://www.datawarehouse.com/resuers/articles/desio.htm>>. Acesso em : 10 jul. 2003.
- DEXTER, M. **Metadata**: data warehouse key. Disponível em: < <http://www.datawarehouse.com/resources/articles/dexter.htm>>. Acesso em: 19 jul.2003
- EIN-DOR, P. **Administração de sistemas de informação**. 2. ed. Rio de Janeiro: Campus, 1985.

FERRANTE, M. V.; RODRIGUEZ, A. J. **Tecnologia da informação e mudança organizacional**. Rio de Janeiro: Infobook, 1995.

FLEURY, A. **Novas tecnologias, capa-citação tecnológica e processo de trabalho**: comparações entre o modelo japonês e o brasileiro. São Paulo: Editora USP, 1993.

FINKELSTEIN, C. **Metadata and XML for business**. Disponível em: <<http://www.tdan.com/i009fe03.htm>>. Acesso em: 12 out. 2003.

GORDON, B. D. **Management information systems**: conceptual foundations, strucre and development. New York: McGraw, 1974.

GRAEML, A. R. **Sistemas de informação**: o alinhamento de tecnologias de informação com a estratégia competitiva. São Paulo: Atlas, 2000.

GUINCHAT, C; MENOU, M. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. corrig. e aum. Brasília. IBICT, 1994.

IKEMATU, R. S. **Gestão de metadados**: sua evolução na tecnologia da informação. Disponível em: <<http://www.pr.gov.br/celepar/celepar/batebyte>>. Acesso em: 11 maio 2003.

KEEN, P. G. W. **Guia gerencial para a tecnologia da informação**. 2. ed. Rio de Janeiro: Campos, 1996.

LEWIS, W. J. Pilar of the community. **Intelligent Enterprise**, San Mateo, v.3, n.13, p. 33-38, Aug. 2000.

LOPES, I. L. Estratégia de busca na recuperação da informação: revisão de literatura. **Ciência da Informação**, Brasília, v. 31, n.2, p. 60-71, maio/ago. 2002.

MARTIN, J. **Engenharia da informação**. Rio de Janeiro: Campus, 1991.

MORAVEC, A. F. Basic concepts for designing a fundamental information system. **Management Services**, v. 2, n. 4, p. 37-45, jul./aug. 1965.

MORESI, E. A. D. Inteligência organizacional: um referencial integrado. **Ciência da Informação**, Brasília, v. 30, n. 2, p. 35-46, maio/ago. 2001.

MOURA, G. A. C. de. **Sistemas de buscas da web**: diretórios e mecanismos de busca. Disponível em: < www.quatrocantos.com.br> Acesso em: 12 out. 2003.

MICHAELLIS: minidicionário escolar da língua portuguesa. São Paulo: Melhoramentos. 2000.

POLLONI, E. G. F. **Administrando sistemas de informação**: estudo de viabilidade. São Paulo: Futura, 2000.

PRETTO, N. de L. **Uma escola sem/com futuro**: educação e multimídia. Campinas: Papirus, 1996.

ROCHA, D. **Noções de HTML**. Disponível em: <<http://www-usr.inf.ufsm.br/~drocha/manual.htm>>. Acesso em: 12 out. 2003.

SANTOS, A. R. **Metodologia científica**: a construção do conhecimento. 5. ed. Rio de Janeiro: DP&A, 2002.

SILVA, H. B. da. Inteligência competitiva na Internet: um processo otimizado por agentes inteligentes. **Ciência da Informação**, Brasília, v. 32, n. 1, p. 115-134, jan/abr. 2003.

SOUZA, T. B. de; CATARINO, M. E.; SANTOS, P. C. dos. Metadados: catalogando dados na Internet. **Transinformação**, v. 9, n. 2, p. 93-105, maio/ago. 1997.

WEIBEL, S. **Dublin Core metadata element set**: reference description. Disponível em: <http://www.purl.org/metadata/dubli_core_elements>. Acesso em: 09 jul. 2003.

APÊNDICE 1 – SITES QUE TRAZEM OS INDICADORES DAS FERRAMENTAS DE BUSCAS MAIS UTILIZADAS NO BRASIL.

A BUSCA. **Genéricos**. Disponível em: < <http://www.abusca.com.br/genericos.shtml>>. Acesso em: 11 out. 2003.

Buscador de buscadores. Disponível em <<http://www.todoesp.es/busca-busca/cgi-bin/neo?Pais=BRASIL>> . Acesso em: 10 out. 2003.

Buscopio. Disponível em: <<http://www.buscopio.net/esp/index.php?cat=213>>. Acesso em: 11 out. 2003.

Guia dos buscadores latinos. Disponível em: <<http://www.latindex.com/buscadores/bra.htm>> . Acesso em: 11 out. 2003.

Mecanismos de busca. Disponível em: <<http://fmail4.uol.com.br/cgi-bin/webmail.exe>>. Acesso em : 10 out. 2003.

MASTER EM JORNALISMO PARA EDITORES. **Guia master de busca e pesquisa**. Disponível em: <<http://www.masteremjornalismo.org.br/master/guia/buscadores.htm>> . Acesso em: 11 out. 2003.

UNIVERSIA BRASIL. **Buscadores**. Disponível em: <<http://www.universiabrasil.net/buscadores/>>. Acesso em 10 out. 2003.

UNIVERSIDADE GUARULHOS. **Buscadores**. Disponível em: <<http://www.ung.br/wpages/especiais/buscadores/buscadores.htm>> . Acesso em: 10 out. 2003.

APÊNDICE 2 - E-MAIL ENVIADO ÀS FERRAMENTAS DE BUSCA: GOOGLE, CADÊ E ALTAVISTA.

Prezados Senhores,

Meu nome é Alexandra Carvalho Guimarães, sou graduanda do último período do Curso de Gestão da Informação da Universidade Federal do Paraná. Meu trabalho de conclusão de curso é a respeito do uso dos metadados nos sites das Instituições de Ensino Superior da rede particular de Curitiba.

Gostaria de esclarecer algumas dúvidas sobre as seguintes questões, a respeito da ferramenta de busca GOOGLE:

Qual a política de indexação e recuperação da ferramenta de busca Google?

A indexação é feita manualmente (diretórios) ou informatizada (índices)?

A ferramenta utiliza o recurso metadado na busca e recuperação das pesquisas?

Muito obrigada pela colaboração.

Atenciosamente,

Alexandra Carvalho Guimarães

APÊNDICE 3 – LISTA DAS INSTITUIÇÕES DE ENSINO SUPERIOR DA REDE PARTICULAR DE CURITIBA E SEUS ENDEREÇOS ELETRÔNICOS.

Centro de Educação Tecnológica Camões – CETESC

Não possui *site*.

Centro de Educação Tecnológica Internacional – CENINTER

www.ceninter.com.br

Centro de Educação Tecnológica Martinus – FELC

www.imec.br

Centro de Educação Tecnológica OPET – CET

www.opet.com.br

Centro Universitário Campos de Andrade – UNIANDRADE

www.uniandrade.br

Centro Universitário Positivo- UNICENP

www.unicenp.br

Escola Superior de Estudos Empresariais e Informática – ESEEI

www.spet.br

Escola Superior de Gestão Comercial e Marketing – ESIC

www.esic.br

Faculdade Anchieta – FAESP

www.faesppr.edu.br

Faculdade Camões

Não possui *site*

Faculdade de Administração, Ciências, Educação e Letras – FACEL

www.facel.com.br

Faculdade de Ciências e Tecnologia do Paraná - FACET-PR

www.facet.br

Faculdade de Ciências Sociais e Aplicadas do Paraná – FACET

www.facet.br

Faculdades Integradas do Brasil – UNIBRASIL

www.unibrasil.com.br

Faculdade de Filosofia São Boaventura
www.saoboaventura.edu.br

Faculdade Dom Bosco – FDB
www.dombosco.com.br

Faculdade Doutor Leocádio José Correia – FALEC
www.falec.br

Faculdade Evangélica do Paraná – FEPAR
<http://www.fepar.edu.br>

Faculdade Evangélica Luterana de Curitiba – FELC
www.imec.br

Faculdade Hoyer de Comunicação Social – FHCS
www.hoyer.edu.br/

Faculdade Internacional de Curitiba – FACINTER
www.facinter.br

Faculdade Organização Paranaense de Ensino Técnico – FAO
www.opet.com.br

Faculdade Padre João Bagozzi - Faculdade Bagozzi
www.bagozzi.com.br

Faculdade Paranaense de Administração – FPA
www.spei.br

Faculdade Pitágoras de Administração Superior de Curitiba – Pitágoras
www.faculdadepitagoras.com.br

Faculdade Teológica Batista do Paraná – FTBP
www.ftbp.com.br

Faculdades Bom Jesus – FBJ
<http://www.fae.edu>

Faculdades Integradas "Espirita" – FIES
www.unibem.br/

Faculdades Integradas Curitiba – FIC
www.faculdaDESCURITIBA.br

Faculdades Integradas Santa Cruz de Curitiba – FARESC
<http://www.santacruz.br>

Faculdades SPEI – FACSPEI
www.spei.br

Instituto de Ciências Sociais do Paraná- ICSP
www.fesppr.br

Instituto de Ensino e Cultura do Paraná – IECP –
Não possui site.

Instituto Sul Brasil de Ensino Superior
<http://www.santacruz.br>

Instituto Superior de Educação Nossa Senhora de Sion - ISE – Sion
www.sion.arauc.br

Instituto Superior de Educação Padre João Bagozzi
www.bagozzi.com.br

Pontifícia Universidade Católica do Paraná – PUC - PR
www.pucpr.br

Unidade de Ensino Superior Expoente – UNIEXP
www.uniexp.edu.br

Universidade Tuiuti do Paraná – UTP
<http://www.utp.br>